Fresenius J Anal Chem (1994) 350:344–351

Fresenius' Journal of Analytical Chemistry

Multivariate control charts: Control charts for calibration curves

Oto Mestek[1], Jiří Pavlík[2], Miloslav Suchánek[1]

[1] Department of Analytical Chemistry, Institute of Chemical Technology, Technická 5, Prague 6, 16628, Czech Republic
[2] Department of Mathematics, Institute of Chemical Technology, Technická 5, Prague 6, 16628, Czech Republic

Received: 16 January 1994 / Accepted: 18 March 1994

Abstract. The proposed multivariate control charts for p-dimensional vectors are an extension of the conventional control charts for one variable. The controlling quantity is the Mahalanobis distance of vector **x** from the central value vector **x**..: $D = (\mathbf{x} - \mathbf{x}..)^T \cdot \hat{\mathbf{C}}^{-1} \cdot (\mathbf{x} - \mathbf{x}..)$, where $\hat{\mathbf{C}}$ is the covariance matrix estimate. The quantity D has Hotelling's T^2 distribution. A PC program was set up for the automatic graphical construction of such charts. The program draws the sequential chart of the quantity D as well as the position of the vectors **x** in the p dimensional control ellipsoid in the axes of the principal components. In this way a control chart was developed for the calibration curve in the photometric determination of Fe^{3+} with sulfosalicylic acid. Vector **x** was formed by absorbance values for the calibration curve points ($p = 5$). The chart can assist in detection of even small disturbances of the calibration curve.

1. Introduction

The control chart method is a useful tool for internal quality control, widely used in analytical and clinical laboratories. The charts are based on the repetitive analysis of a sample (or samples) of known composition (daily, with each batch, etc.) and the graphical representation of the quality indices examined, such as the results of individual analyses, averages of groups of analyses, or the ranges within groups. Plots of such indices against time can be used to examine the stability of the system or to draw conclusions concerning the nature of the variability of the results. It is clear that factors affecting variability include, in addition to those intrinsic to the method, also the quality of the laboratory equipment, chemicals, personnel skills, etc.

Shewhart charts are probably the best-known control charts [1]. They are set up by drawing a central line and a pair of lines of control limits, parallel to the central line above and below it; the limits are usually referred to as the inner and outer control limits.

Correspondence to: O. Mestek

2. Theoretical

2.1 Control charts for one variable

There are no problems in constructing the chart if a single type of data concerning the amount of a single analyte is treated. In the first stage, when the method is stable, several analyses of a reference material are made every day or with each sample batch (an intralaboratory standard is sufficient for such purpose). For statistical reasons, the number of groups of analyses should not be lower than 20, and the analyses should not all be accomplished within a short period of time because data concerning the significance of variability among the groups are not known in advance [2]. Data so obtained are then used to determine the limit of "tolerable error" i.e., they will serve as a reference set for future control analyses of the same reference material. The analysis of the data variance in the starting series then enables several kinds of control charts to be constructed. The first is the control chart for individual measurements; here the central line will have a value which is equal to the mean value of all measurements and the inner and outer control limits will lie at distances of $\pm 2 \cdot s_T$ and $\pm 3 \cdot s_T$, respectively, from the central line. (s_T is the total standard deviation).

$$s_T^2 = \frac{\sum_{i=1}^{q} \sum_{j=1}^{r} (x_{ij} - x..)^2}{q \cdot r - 1} \quad (1)$$

The group-average chart is the next control chart. The central line again represents the average value of all measurements and the control limits lie at distances of $\pm 2 \cdot s_S$ and $\pm 3 \cdot s_S$ from it, where s_S is the standard between-group deviation.

$$s_S^2 = \frac{\sum_{i=1}^{q} (x_i. - x..)^2}{q - 1} \quad (2)$$

In both cases, q has the meaning of the number of groups and r is the number of repetitions within each group.

The coefficients of ± 2 and ± 3 for the standard deviations were initially ± 1.96 and ± 3.09, for which on average 5% of all results lie beyond the inner control limits and 0.2% of all results lie beyond the outer control limits. For practical reasons the coefficients are rounded to the integers ± 2 and ± 3, so that 4.55% and 0.27% results lie beyond the inner and outer control limits, respectively.

The Shewhart control charts are usually interpreted as follows. An adjustment is made if:

- one value lies beyond the outer control limits,
- two values lie beyond the inner control limits on one side of the chart,
- a sequence of 7 points lies on one side of the control chart (bias),
- a sequence of 7 points exhibits an increasing or decreasing trend.

2.2 Control charts for several variables

The use of one dimensional statistical methods is questionable if a vector **x**, expressing the amounts of several analytes rather than data concerning the amount of a single analyte, emerges from the analysis. Sets of one-dimensional control charts are inapplicable, especially when the components of vector **x** are mutually correlatable. There is a principle saying that what emerged together should be analyzed together. The multivariate control chart proposed is designed to treat such data.

If p-dimensional vectors $\mathbf{x}_i$ ($i = 1, \ldots, n$) have a p-dimensional normal distribution with a mean value of μ and a covariance matrix **C**, then the Mahalanobis distance of the vectors from the mean value,

$$D_i = (\mathbf{x}_i - \boldsymbol{\mu})^T \cdot \mathbf{C}^{-1} \cdot (\mathbf{x}_i - \boldsymbol{\mu}) \tag{3}$$

has the χ^2 distribution with p degrees of freedom (see, e.g., [3], [5], [6] – also for the basic statistical methods to be used later). This fact is utilized in the construction of the control chart of vectors $\mathbf{x}_i$. For $100 \cdot \alpha\%$ of vectors $\mathbf{x}_i$, Eq. 3(a)

$$(\mathbf{x}_i - \boldsymbol{\mu})^T \cdot \mathbf{C}^{-1} \cdot (\mathbf{x}_i - \boldsymbol{\mu}) \leq \chi^2_\alpha(p), \tag{3a}$$

is an inequality that describes a p-dimensional ellipsoid. For $\alpha = 0.9973$, which corresponds to the control limit of $\pm 3 \cdot s$ in the one-dimensional charts, 0.27% of the vectors $\mathbf{x}_i$ lie beyond the ellipsoid, whereas for $\alpha = 0.9545$, corresponding to the limit of $\pm 2 \cdot s$, 4.55% of the vectors $\mathbf{x}_i$ lie beyond the ellipsoid.

In addition to the control chart for the quantity D_i, the vectors **x**, transformed into the principal component coordinates, can also be plotted in a control ellipsoid. Inequality (3a) can be modified [4] to the form

$$\sum_{j=1}^{p} y_j^2/\lambda_j \leq c \tag{4}$$

where y_j is the value of the j-th coordinate of the vector in the principal component coordinate system and λ_j is the j-th eigenvalue of the covariance matrix. This is the equation of a p-dimensional ellipsoid whose axes coincide with those of the principal components; the ellipsoid contains all vectors whose Mahalanobis distance from the mean value does not exceed c. If c is put equal to $\chi^2_{0.9545}(p)$, an inner control ellipsoid exists whereas if c is put equal to $\chi^2_{0.9973}(p)$, an outer control ellipsoid is produced.

In the common case, when the true values of $\boldsymbol{\mu}$ and **C** are not known, the Mahalanobis distance must be calculated using estimates:

$$D_i = (\mathbf{x}_i - \mathbf{x}.)^T \cdot \hat{\mathbf{C}}^{-1} \cdot (\mathbf{x}_i - \mathbf{x}.) \tag{5}$$

and the χ^2 distribution must be replaced by Hotelling's T^2 distribution, which can be enumerated using F distribution:

$$T^2(p, n-p) = F(p, n-p) \cdot p \cdot (n-1)/(n-p) \tag{6}$$

This situation differs from those in the case of a univariate control chart, where, for $n > 20$, the normal distribution is very close to the more exact Student's t distribution.

The construction of the chart again begins by collecting $\mathbf{x}_{ij}$ data for an initial series of analyses of a reference material, distributed among q groups with r repetitions. As regards the time sequence of data collection, the same principle applies as with one dimensional control charts. The mean value vector $\mathbf{x}..$ is calculated:

$$\mathbf{x}.. = \frac{\sum_{i=1}^{q} \sum_{j=1}^{r} \mathbf{x}_{ij}}{q \cdot r} \tag{7}$$

The value of this vector is subtracted from all vectors $\mathbf{x}_{ij}$. The vectors $\mathbf{x}_{ij} - \mathbf{x}..$ are arranged in a matrix $\mathbf{X}_{Centr}$ whose dimension is $p \times q \cdot r$, and the estimate of the total covariance matrix is calculated:

$$\hat{\mathbf{C}}_T = (\mathbf{X}_{Centr.} \cdot \mathbf{X}^T_{Centr})/(q \cdot r - 1) \tag{8}$$

For the vectors **x** obtained subsequently, the D values are calculated from Eq. (5) using matrix $\hat{\mathbf{C}}_T$ and plotted in the control chart (see later). This control chart is analogous to that for individual measurements and involves a variability both within the groups and between them. To construct the chart of group averages $\mathbf{x}_i.$, those in the initial series are first calculated:

$$\mathbf{x}_i. = \frac{\sum_{j=1}^{r} \mathbf{x}_{ij}}{r} \quad i = 1..q \tag{9}$$

After centering, by subtracting the vector $\mathbf{x}..$, the group averages are arranged into a $p \times q$ matrix ($\bar{\mathbf{X}}_{Centr}$) and an estimate of the covariance matrix of the group averages is calculated:

$$\hat{\mathbf{C}}_S = (\bar{\mathbf{X}}_{Centr} \cdot \bar{\mathbf{X}}^T_{Centr})/(q-1) \tag{10}$$

For the resulting vectors of group averages of analyses of reference samples, the D values are calculated from Eq. (5) using matrix $\hat{\mathbf{C}}_S$ and are drawn in the control chart.

In contrast to the one dimensional case, the resulting control chart is not symmetric and is made up of three parallel straight lines. The first line corresponds to D = 0, (i.e. a zero deviation of vector $\mathbf{x}_{ij}$ or $\mathbf{x}_i$. from vector $\mathbf{x}..$); this is analogous to the central line in the one dimensional chart.

The inner control limit lies at a distance of $T^2_{0.9545}$ from the base line, which is analogous to the control limit of $\pm 2 \cdot s$. The outer control limit lies at a distance of $T^2_{0.9973}$, analogous to the control limit of $\pm 3 \cdot s$. Thus only the magnitude of the deviation of the vector from the central value can be inferred from the chart. To enable the direction of the deviation to be also followed, it is necessary to draw the individual $\mathbf{x}_{ij}$ or $\mathbf{x}_i$, points in the p dimensional control ellipsoid constructed according to Eq. (4). Unfortunately there is no way to draw this control ellipsoid in a simple manner; it is carried out by resorting to sections in the axes of the coordinate pair. There are $p \cdot (p-1)/2$ such sections. The radii of the control ellipsoid in the axes of the principal components are $\sqrt{(T^2_{0.9545} \cdot \lambda_i)}$ for the inner control limit and $\sqrt{(T^2_{0.9973} \cdot \lambda_i)}$ for the outer control limit. Although the relationship of the principal components to the initial variables x may not be interpreted easily, the charts illustrate the trends. It is, of course, also possible to draw projections of the p dimensional control ellipsoid, which is generally rotated around the coordinates of the initial variables x, on to the planes formed by pairs of the initial variables x.

3. Practical

3.1 Computer program for constructing multivariate control charts

To facilitate and automate the construction of multivariate control charts, a PC-program has been developed which is capable of handling up to six variables, measured in groups with up to six repetitions. After reading data of the initial series of measurements, the program tests the normality of distribution of the variables using the Kolmogorov test and the normality of the total p dimensional distribution using the multidimensional skewness test. Normality of the distribution is a necessary assumption for the successful application of the method. Subsequently the covariance matrices are calculated and the ANOVA is performed. The D values of the vectors in the initial series can also be calculated according to Eq. (5) (the series is mapped onto itself), so that the quality of the series is tested.

For the points tested, the program calculates the group average, the D value for that average and its significance. The position of the tested point within the multidimensional control ellipsoid, drawn in the principal component axes, is another important value. This position is expressed by the number labelling the section of the ellipsoid within which the point lies. These and other data are entered into a table, which can be displayed on the screen. The fraction of variability born by the principal component and its composition from the initial variables can be easily seen because the program also performs the PCA for the normalized initial variables.

The program then graphically draws the positions of the tested vectors in the multidimensional control ellipsoid. Sections of the ellipsoid are drawn along axes of any arbitrary combination of the principal component pairs. However one drawback of this approach is that, during the projection into the plane, those points actually lying beyond the p-dimensional ellipsoid may also be projected into the section through the ellipsoid; in fact, they may be projected into all the possible sections. The projection of the control ellipsoid onto planes formed by arbitrary pairs of the initial variables is also possible.

The one dimensional control charts of the individual principal components with the control limits of $\pm 2 \cdot \sqrt{(\lambda_i)}$ and $\pm 3\sqrt{(\lambda_i)}$ and the one dimensional control charts with the limits of $\pm 2 \cdot s_i$ and $\pm 3 \cdot s_i$ for the individual initial variables can also be drawn. The charts, the latter in particular, are only auxiliary because even if all the variables (components of vector **x**) lie within their respective control limits, this does not guarantee that vector **x** lies within the control limits of T^2, (and *vice versa*).

All charts are set up for averages of groups of measurements; it is, however, possible to look into a group and test the positions of the points which make up the group.

3.2 Testing the calibration curve stability

Multivariate control charts can also be utilized in ways other than for the quality control of the entire analytical process. They can also be employed to check the stability of partial analytical operations (e.g. in calibration curve plotting).

3.2.1 Method tested. The multivariate control chart was applied to the particular case of testing the stability of the calibration curve in the photometric determination of Fe^{3+} with sulfosalicylic acid. The calibration curve was set up for the following procedure. Volumes of 0, 1, 2, 3 and 4 ml of a stock solution of 50 μg/ml Fe^{3+} (intralaboratory standard) were added to 50 ml volumetric flasks and diluted with water to 25 ml; 2.5 ml of a 20% solution of sulfosalicylic acid and 1.5 ml of a concentrated solution of ammonia were added; the whole was made up to the mark with distilled water and homogenized. The absorbances of the solutions at 420 nm were measured on a Spekol 11 (Carl Zeiss, Jena) using 1 cm cells.

All points of the calibration curve were measured in duplicate except for the blank, which was prepared in triplicate; one point, chosen at random, served as the reference solution in the photometric measurement.

3.2.2 Data measured. Repeated measurements under independent conditions were performed to obtain 22 calibration curves (q = 22) with duplicate measurements (r = 2) of absorbances of points of 0, 50, 100, 150 and 200 μg Fe^{3+}; this gave five-dimensional vectors **x**(p = 5). The absorbances (multiplied by 10^3) are given in Table 1. The data measured can be characterized by the total

Table 1. Absorbances measured for the initial group of measurements

No.	Blank	50 μg Fe^{3+}	100 μg Fe^{3+}	150 μg Fe^{3+}	200 μg Fe^{3+}
		Absorbance × 10^3			
1	1	104	206	307	409
	3	104	206	308	412
2	4	104	206	308	412
	2	103	204	307	413
3	3	105	207	311	414
	2	104	207	309	411
4	4	104	206	308	411
	2	104	207	312	413
5	−9	92	195	296	397
	−8	95	197	299	400
6	3	107	209	311	412
	3	105	207	308	410
7	3	104	207	311	414
	2	105	208	308	410
8	2	105	208	310	412
	2	104	208	309	412
9	−6	95	196	297	401
	−7	94	197	300	401
10	2	104	206	311	413
	4	105	207	310	412
11	1	103	205	309	412
	2	104	206	307	411
12	−7	94	198	298	404
	−7	96	199	301	402
13	5	105	210	313	415
	7	107	208	315	415
14	3	106	208	311	411
	2	104	207	308	414
15	−8	94	196	299	400
	−6	95	199	302	404
16	4	104	207	311	415
	6	106	210	310	413
17	2	105	206	308	410
	4	106	208	310	413
18	2	104	206	309	414
	0	103	206	308	409
19	0	101	203	305	409
	1	102	206	307	411
20	1	104	206	311	410
	4	106	208	309	414
21	−9	92	194	298	400
	−10	92	194	297	398
22	−8	95	195	298	401
	−8	95	199	301	403

covariance matrix $\hat{C}_T$ (Eq. (8)) and the covariance matrix of group averages $\hat{C}_S$ (Eq. (10)), which are as follows:

$$\hat{C}_T = \begin{bmatrix} 24.27 & 23.23 & 23.21 & 23.96 & 25.39 \\ 23.23 & 23.31 & 23.08 & 23.64 & 24.47 \\ 23.21 & 23.08 & 23.85 & 23.97 & 24.83 \\ 23.96 & 23.64 & 23.97 & 26.30 & 26.17 \\ 25.39 & 24.47 & 24.83 & 26.17 & 28.72 \end{bmatrix}$$

and

$$\hat{C}_S = \begin{bmatrix} 24.22 & 23.40 & 23.44 & 24.40 & 24.85 \\ 23.40 & 23.34 & 23.25 & 23.87 & 24.85 \\ 23.44 & 23.25 & 23.54 & 24.06 & 25.03 \\ 20.40 & 23.87 & 24.06 & 25.51 & 26.08 \\ 25.52 & 24.85 & 25.03 & 26.08 & 27.63 \end{bmatrix}$$

The coordinates of the vector of the mean values are [− 0.2, 101.6, 204.0, 306.5, 409.0]. (All data are absorbances multiplied by a factor of 10^3.)

Visual comparison of the two matrices reveals that the inter-group variability will be more significant than the residual (intra-group) variability. This is confirmed by a multidimensional variance analysis using the Bartlett approximation (which is included in the computer code): the quantity Θ is calculated:

$$\Theta = \frac{|(q \cdot r - 1) \cdot \hat{C}_T - (q - 1) \cdot r \cdot \hat{C}_S|}{|(q \cdot r - 1) \cdot \hat{C}_T|} \tag{11}$$

Its function b,

$$b = [1 + (p + q)/2 - q \cdot r] \cdot \ln(\Theta) \tag{12}$$

has an χ^2 distribution with $p \cdot (q - 1)$ degrees of freedom. The b value obtained was 196.10, corresponding to a confidence level higher than 0.999. The higher value of the

intergroup variability can be explained in terms of the instrumental stability of the Spekol instrument, which is a simple apparatus designed for routine applications.

The components of the vectors **x** are highly mutually correlated, as demonstrated by the correlation matrix of group averages $\hat{\mathbf{R}}_S$:

$$\hat{\mathbf{R}}_S = \begin{bmatrix} 1 & 0.98 & 0.98 & 0.98 & 0.99 \\ 0.98 & 1 & 0.99 & 0.98 & 0.98 \\ 0.98 & 0.99 & 1 & 0.98 & 0.98 \\ 0.98 & 0.98 & 0.98 & 1 & 0.98 \\ 0.99 & 0.98 & 0.98 & 0.98 & 1 \end{bmatrix}$$

Consistent with this is the fraction of total variability, explained by the first principal component, which is 98.6%. The loadings (the correlation coefficients between the principal component and the initial component) of the first principal component are 0.99, 0.99, 0.99, 0.99, and 0.99, respectively. The loadings of the remaining principal components are invariably lower than 0.11 in their absolute value.

3.2.3 Control chart of the calibration curve. The 22 calibration curves obtained were used as the starting series for setting up a control chart. The tested vector **x** was directly formed by absorbances of the calibration curve points. The mapping of the series onto itself is apparent from Fig. 1 showing the sequential diagram of the Mahalanobis distance D. Figure 2 shows a section through the control ellipsoid in the axes of the first and second principal components. The distribution of the points into two groups according to the value of the first principal component is apparent; this distribution also corresponds to that into groups with respect to the displacement along the absorbance axis. The uniformity of loadings of the individual variables for the first principal component indicates that this component is a measure of the calibration curve displacement along the absorbance axis; any disturbance of the correlation structure (linearity of the curve) is only mirrored by the following principal components. Thus the distribution of points in Fig. 3 (control ellipse in the axes of the 3rd and 4th principal components) is uniform. Figure 4 shows the sequential control chart for

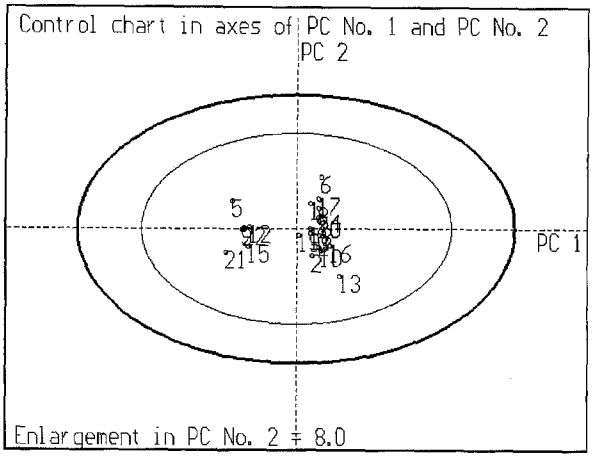

Fig. 2. Section through the control ellipsoid in the axes of the 1st and 2nd principal components for the starting series of measurements

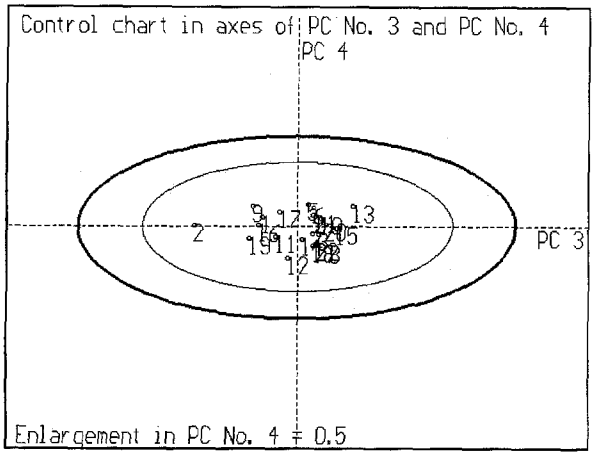

Fig. 3. Section through the control ellipsoid in the axes of the 3rd and 4th principal components for the starting series of measurements

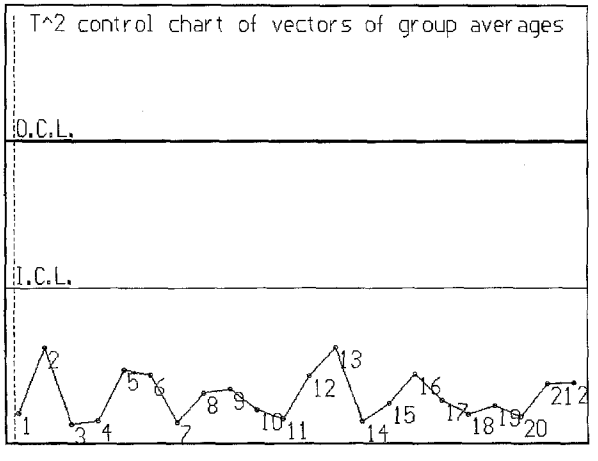

Fig. 1. Control diagram of the Mahalanobis distance for the starting series of measurements

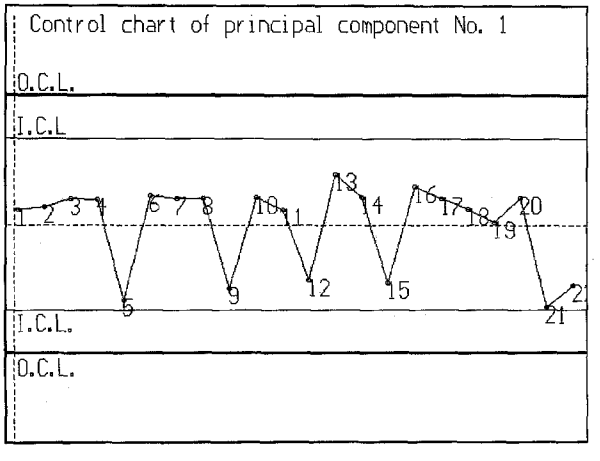

Fig. 4. Control chart of the 1st principal component for the starting series of measurements

the first principal component. The distribution of points again corresponds to the different displacement of the calibration curves along the absorbance axis. Point No. 21 lies near the inner control limit; however the control limits in one dimensional charts (where a variable is considered on its own, independent of the others) are different from those in a control ellipsoid (where each variable is considered in relation to the remaining variables). The sensitivity of the control chart for indicating various disturbances of the calibration curve was tested on an additional 13 calibration curves into which the following errors had been deliberately introduced:

1 – correct calibration curve,
2 – measured at 410 nm instead of 420 nm,
3 – measured at 430 nm instead of 420 nm,
4 – a double amount of sulfosalicylic acid added to one point of the calibration curve,
5 – a double amount of ammonia added to one point of the calibration curve,
6 – the blank reference solution actually contained 5 μg Fe^{3+}
7 – one of two repetitions of the blank solution actually contained 5 μg Fe^{3+},
8 – one of two repeated measurements of the point of 50 μg Fe^{3+} contained an additional 5 μg Fe^{3+},
9 – one of two repeated measurements of the point of 100 μg Fe^{3+} contained an additional 5 μg Fe^{3+},
10 – one of two repeated measurements of the point of 150 μg Fe^{3+} contained an additional 5 μg Fe^{3+},
11 – one of two repeated measurements of the point of 200 μg Fe^{3+} contained an additional 5 μg Fe^{3+},
12 – stock solution of Fe^{3+} contained 51 μg/ml Fe^{3+} instead of 50 μg/ml Fe^{3+},
13 – correct calibration curves measured in 60 min.

The absorbances measured are given in Table 2.

Figure 5 demonstrates that the calibration curve is relatively insensitive to the amounts of reagents added, to small wavelength shifts towards higher values and to a delay in measurement; appreciable effects are caused by wavelength shifts towards lower values, addition of small amounts of Fe^{3+} and changes in slope of the calibration curve. Points 2 and 8–12 lie beyond the outer control limit, points 6 and 7 lie beyond the inner control limit. Figure 6 shows that only the disturbance of point 6 (reference solution containing additional Fe^{3+} – displacement

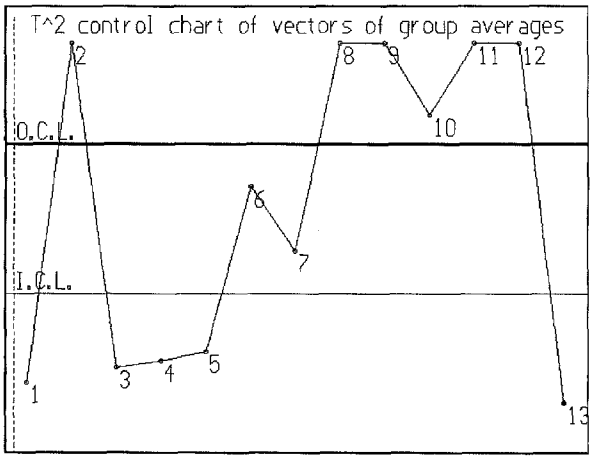

Fig. 5. Control chart of the Mahalanobis distance for the tested group

Table 2. Absorbances measured for the group of tested measurements

No.	Blank	50 μg Fe^{3+}	100 μg Fe^{3+}	150 μg Fe^{3+}	200 μg Fe^{3+}
			Absorbance × 10^3		
1	1	103	204	306	407
	1	104	205	309	411
2	4	98	195	292	389
	2	99	195	294	393
3	1	102	204	308	411
	2	103	204	304	406
4	− 10	91	195	295	397
	− 10	90	193	299	400
5	− 10	91	193	295	397
	− 9	90	194	296	401
6	− 20	84	184	285	391
	− 21	82	185	289	389
7	10	106	208	310	414
	3	104	208	312	412
8	3	117	207	312	414
	2	108	207	309	413
9	2	105	217	309	411
	1	103	206	310	412
10	− 9	92	196	297	398
	− 9	92	194	306	399
11	− 11	93	196	299	411
	− 10	94	195	298	401
12	2	109	212	318	424
	2	109	212	317	420
13	− 1	102	205	305	410
	0	102	204	309	411

along the absorbance axis) affects the value of the first principal component; the other disturbances are only mirrored by subsequent principal components (Fig. 7). Figure 8 is an example of projection of the control ellipsoid onto the plane formed by the initial variables; in this case the "blank" and the "50 μg" variables.

3.2.4 Linear regression. The absorbances in the initial series of measurements were fitted by regression straight lines,

$$\hat{A} = a_0 + a_1 \cdot m \tag{13}$$

where $\hat{A}$ is the observed absorbance estimate, m is the Fe^{3+} content in the solution and a_0 and a_1 are constants calculated by the least squares method. The values of the constants are given in Table 3. The quantity s_y in the table is the square root of the residual variance of the dependence:

$$s_y^2 = \sum_{i=1}^{q} \sum_{j=1}^{r} (A_{ij} - a_0 - a_1 \cdot m_{ij})^2/(n-2) \tag{14}$$

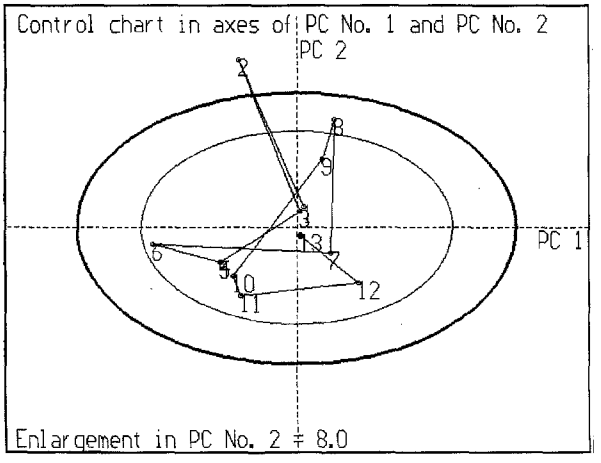

Fig. 6. Section through the control ellipsoid in the axes of the 1st and 2nd principal components for the tested group

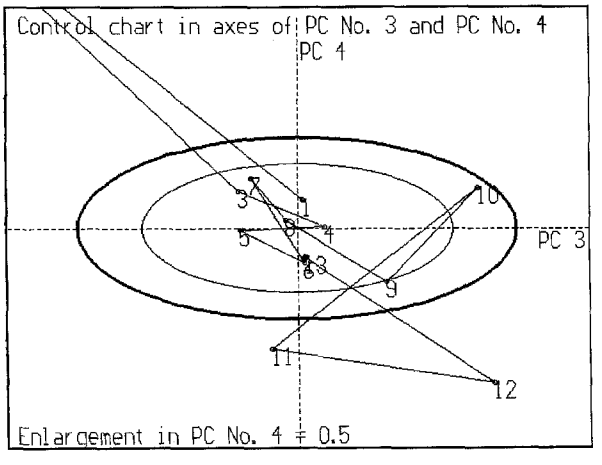

Fig. 7. Section through the control ellipsoid in the axes of the 3rd and 4th principal components for the tested group

where q is the number of points in the calibration curve (q = 5), r is the number of measurements in each point (r = 2), n is the total number of measurements (n = q · r = 10) and A is the true absorbance value. The quantity F in the table is a measure of the linearity of the dependence:

$$F = \frac{n-q}{q-2} \cdot \frac{\sum_{i=1}^{q} r \cdot (A_{i.} - \hat{A}_i)^2}{\sum_{i=1}^{q} \sum_{j=1}^{r} (A_{ij} - A_{i.})^2} \tag{15}$$

The calculated F value is invariably lower than $F_{0.95}(3, 5) = 5.410$; hence all of the dependences are really linear. The highest slope, 2.052, is found for dependence 18, whereas the lowest slope, 2.036, is observed for dependence 5. Both curves were tested for a parallel course.

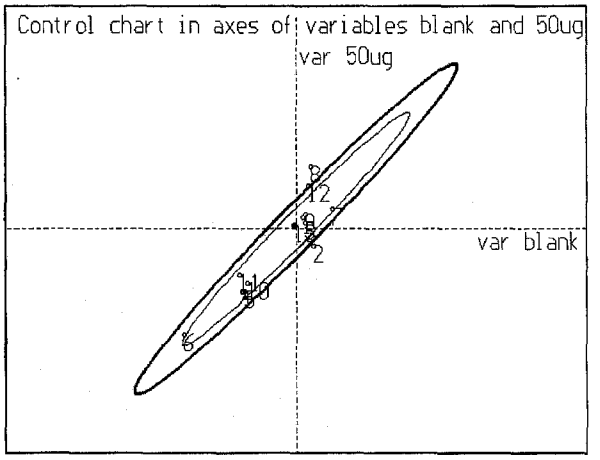

Fig. 8. Projection of the control ellipsoid into a plane formed by initial blank variables and 50 μg Fe for the tested group

Table 3. Linear regression for the initial group of measurements

No.	a_0	a_1	s_y	F
1	1.90	2.041	0.9969	0.226
2	1.70	2.046	1.5930	4.485
3	2.20	2.051	1.0031	0.122
4	2.30	2.048	1.4009	0.427
5	− 8.20	2.036	1.4832	0.167
6	3.60	2.039	1.2475	0.310
7	2.40	2.048	1.3416	0.048
8	2.20	2.050	0.5701	2.667
9	− 7.00	2.038	0.9618	0.389
10	2.40	2.050	0.9618	1.417
11	1.10	2.049	0.8624	0.813
12	− 7.10	2.049	1.2016	0.472
13	4.80	2.052	1.6125	2.667
14	2.50	2.049	1.2425	0.049
15	− 7.30	2.048	1.6317	0.154
16	3.90	2.047	1.4979	1.053
17	3.10	2.041	1.2016	0.083
18	0.90	2.052	1.4009	0.022
19	− 0.20	2.047	1.3252	0.798
20	2.50	2.048	1.5370	0.036
21	− 9.90	2.045	0.8023	1.194
22	− 7.80	2.049	1.3624	0.040

The criterion tested,

$$t = \frac{|a_1 - a_1'|}{s} \tag{16}$$

$$s^2 = \frac{s_y^2 + s_y^{2\prime}}{2} \cdot \frac{2}{\sum_{i=1}^{n} (x_i - \bar{x}.)^2} \tag{17}$$

has the Student distribution with $2 \cdot n - 4$ degrees of freedom. The calculated value of 1.75 is lower than $t_{0.975}(16) = 2.120$ (both-sided test); hence, the two straight lines are parallel, and it is reasonable to assume that the remaining straight lines of the initial series will also be parallel.

The regression dependences were also calculated for the group of calibration curves with disturbances deliberately introduced. The results are given in Table 4. All the dependences are linear. Since straight line 1 corresponds to the correct calibration, it served as the reference with which all the remaining dependences were compared. The comparison was based on the parallel course test (Eqs. (16) and (17)) and on the F test to compare the variances s_y^2. The results of the test and the comparison with the test performed by means of the control chart are given in Table 5, where detection of a disturbance is labelled with the symbol !. The critical values are $t_{0.975}(16) = 2.120$ and $F_{0.95}(8,8) = 3.438$. The table demonstrates that the disturbance for the calibration dependence 6 (incorrect reference solution – displacement along the absorbance axis) is only detected by the control chart; however, in view of the high variance of displacements along the absorbance axis for the individual dependences, this may be fortuitous: If the displacement due to the incorrect reference solution did not have the same direction as that caused by the instrument, the control chart would also fail to detect the disturbance. Thus the control chart revealed all disturbances disclosed by the regression analysis and, in addition, signalled a disturbance for point 7 which regression analysis failed to detect. Although not appreciably different from that of the correct straight line, the variance s_y^2 for straight line 7 at the 0.95 confidence level is higher ($F = 2.95$). Hence, in this case of testing the stability of the calibration curve, the control chart is more sensitive to disturbances than regression analysis.

Table 4. Linear regression for the group of tested measurements

No.	a_0	a_1	s_y	F
1	1.10	2.040	1.3647	0.173
2	2.00	1.941	1.5350	0.847
3	1.00	2.035	1.7048	0.136
4	− 10.70	2.047	1.5632	0.506
5	− 10.60	2.044	1.4491	1.133
6	− 20.20	2.050	1.3509	0.205
7	4.90	2.038	2.3425	0.732
8	5.20	2.040	3.9937	2.957
9	2.50	2.051	3.5768	0.999
10	− 9.30	2.049	3.1295	1.370
11	− 11.00	2.076	2.9496	0.564
12	2.80	2.097	1.3252	1.088
13	− 0.70	2.054	1.1347	0.140

Table 5. Comparison of the control charts and regression analysis. The symbol ! denotes detection of a disturbance

No	t	F	Cont. charts
2	10.78 !	1.26	!
3	0.51	1.56	
4	0.75	1.31	
5	0.45	1.13	
6	1.16	0.98	!
7	0.19	2.95	!
8	0.00	8.56 !	!
9	0.64	6.83 !	!
10	0.59	5.26 !	!
11	2.48 !	4.67 !	!
12	6.70 !	0.96	!
13	1.76	0.69	

4 Conclusions

The multivariate control charts proposed are based on, and are an extension of, the conventional control charts for one variable. They can be applied to the quality control of analytical procedures whose output consists of information on the amounts of several analytes and during the control of partial operations of the individual analytical methods. The control chart for the calibration curve of the photometric determination of Fe^{3+}, demonstrated above, is an example of such an application. The procedure is very sensitive to disturbances in the correlation structure (linearity) of the curves but is less sensitive to displacements of the curves along the absorbance axis. This, however, is due to the high scatter in the displacements (which is a consequence of insufficient stabilization of the measuring instrument) rather than to the control chart itself.

The PC-program for the construction of the charts is available from the authors.

References

1. Kateman G, Pijpers FW (1981) Quality control in analytical chemistry. Wiley, New York, pp 107–110
2. Massart DL, Vandeginste BGM, Deming SN, Michotte Y, Kaufman L (1988) Chemometrics: a textbook. Elsevier, Amsterdam, pp 93–96
3. Lehmann LE (1959) Testing statistical hypotheses. Wiley, New York
4. Bolch BW, Huang CJ (1974) Multivariate statistical methods for business and economics. Prentice Hall, New York. Russian translation (1979) Mnogomernyje statisticeskije metody dlja ekonomiki. Statistika, Moscow, pp 25–28
5. Anderson TW (1958) An introduction to multivariate statistical analysis. Wiley, New York
6. Rao RC (1965) Linear Statistical Interference and its Applications, Wiley, New York